Received November 9, 2017, accepted December 22, 2017, date of publication January 1, 2018,
date of current version February 28, 2018.

Digital Object Identifier 10.1109/ACCESS.2017.2788918

Analyzing Technological Spillover Effects Between Technology Classes: the Case of Korea Technology Finance Corporation

SUNGCHUL CHOI[1], MAENG SEOK NOH[2], JANGHYEOK YOON[3], HYUNSEOK PARK[4], AND WONCHUL SEO[5]

[1]Department of Industrial and Management Engineering, Gachon University, Seongnam 13120, South Korea
[2]Department of Statistics, Pukyong National University, Busan 48513, South Korea
[3]Department of Industrial Engineering, Konkuk University, Seoul 05029, South Korea
[4]Department of Information System, Hanyang University, Seoul 04763, South Korea
[5]Graduate School of Management of Technology, Pukyong National University, Busan 48513, South Korea

Corresponding author: Wonchul Seo (wcseo@pknu.ac.kr)

This work was supported by the National Research Foundation of Korea through the Basic Science Research Program funded by the Ministry of Science, ICT and Future Planning under Grant NRF-2014R1A1A1005317 and Grant NRF-2015R1C1A1A01056185.

ABSTRACT A technology evaluation system is mandatory to successfully implement a technology-based financial support system. Technology evaluation has generally been relied on the experts' manual work. Various quantitative indicators have been presented to improve the efficiency of this manual work. Among these indicators, the spillover effect is perceived as useful for the disposal of patents of a firm, which received credit guarantee but lost its ability to service its debt. A model for measuring the spillover effects has already been proposed, but it has low reliability. Therefore, this paper presents a systematic approach for measuring technological spillover effects between technology classes. The approach mainly relies on patent data due to its features of the latest reliable sources of technological intelligence. We first extract co-classification information from patent data and generate association rules between technology classes. The relationships represented by the rules, however, can only depict the direct effects. Therefore, we first derive the indirect effects from the direct ones and then integrate both the effects to measure the technological spillover effects. We conduct an empirical study to show the applicability of the presented approach using patents granted in the Korean Intellectual Property Office. We expect that this paper can contribute to establish a quantitative evaluation model to help assess technologies for successful technology-based credit guarantee system. It will improve the reliability of the technology assessment by reducing the variance of the qualitative evaluation results due to the individual differences of the evaluator. Furthermore, it will also enhance the efficiency of evaluation work.

INDEX TERMS Association rule mining, DEMATEL, technological spillover effect, technology evaluation, technology financing.

I. INTRODUCTION

Technology financing as a technology-based financial support system for high-tech firms has been crucial to secure a sustainable growth power since it can help overcome the problem of lack of financial resources arose by the conventional collateral-based lending practice of banks [1], [2]. One possible solution for the realization of technology financing might be to provide credit guarantees to firms with technology [3]. It regards the technology as collateral and provides financial support that corresponds to the value of the collateral. Therefore, to successfully implement the technology-based credit guarantees, a system that can clearly assess the value of technology is required. Korea Technology Finance Corporation (KOTEC) is the most representative non-profit credit guarantee institution [4]. KOTEC has developed a Kibo Technology Rating System (KTRS) that includes 33 detailed indicators to assess the potential future value of technology. The KTRS conducts an assessment of the technology in both quantitative and qualitative aspects. The quantitative evaluation is performed mainly in technology classes and the qualitative evaluation is done in individual technology.

Among the detailed indicators in the KTRS, the technology spillover effect determines how much the knowledge implied by the technology affects the entire industry, and ultimately represents the ability to derive new inventions [5]. The spillover effect is naturally recognized as an important criterion because it is useful for the disposal of patents (or strictly speaking patented technologies) of a firm with a credit guarantee. When the firm loses its ability to service its debt, KOTEC will dispose of its technologies through technology transfer or sale. In this case, technologies with high spillover effects are generally considered to be easy to dispose of. The current evaluation model for the spillover effects in the KTRS uses patent citation information and odds ratio on a contingency table, but it has low reliability. Therefore, this study presents a systematic approach for measuring technological spillover effects between technology classes. The approach mainly relies on patent data due to its features of the latest reliable sources of technological intelligence [6], [7]. Patent data are widely incorporated into an analysis of technological trends including examining important factors for the invention of biology-related technologies [8], identifying potential opportunities for new products or technology development [7], [9], analyzing technology trends [10], and investigating knowledge spillovers [11].

The presented approach first extracts co-classification information from patent data. A patent is naturally classified into multiple International Patent Classification (IPC) codes. The co-classification information can be generated by extracting IPCs assigned together in each patent. And then, association rule mining is applied to obtain connection relationships between IPCs using the extracted information. Association rule mining discloses interesting relationships among various items examining their co-occurrences in a dataset [12], [13]. The relationships, however, only depict direct connections between technology classes so our approach also utilizes a Decision Making Trial and Evaluation Laboratory (DEMATEL) method to generate comprehensive spillover effects including direct and indirect influential relationships. DEMATEL as one of network analysis-based decision making techniques aims to examine elements in complicated systems and create meaningful relationships between elements by exploring the extent that each element exerts on others [14]–[16]. To show the applicability of the presented approach, we carry out a case study using patents granted in the Korean Intellectual Property Office (KIPO). Moreover, to examine the feasibility of the approach, we investigate how the evaluation results of the approach explain the probability of default of the firms who have been guaranteed by the KOTEC. The biggest risk factor for the KOTEC is that the guaranteed firm is in default because the KOTEC has the obligation to pay off its debt. This is why we use the probability of default in the feasibility verification process. To do that, we build a logit model between the extent of the spillover effects and the default probability of each technology class. And then, we conduct a Receiver Operating Characteristic (ROC) curve analysis to determine the extent to which the approach explains the default probability by calculating the Area Under the ROC curve (AUC) which indicates a measure of the detection capability [17]. We expect that this study can contribute to establish a quantitative evaluation model to help assess technologies for successful technology-based credit guarantee system. It will improve the reliability of the technology assessment by reducing the variance of the qualitative evaluation results due to the individual differences of the evaluator. Furthermore, it will also contribute to enhance the efficiency of evaluation work.

II. GROUNDWORK

A. TECHNOLOGICAL SPILLOVER EFFECT ANALYSIS

Patents represent the trends of technological innovation and development as a result of inventions with high technological reliability [18], [19]. The patent analysis enables us to illustrate the technological knowledge flow that occurs between various technological classes, and furthermore, to quantify the extent of the technological spillover effects based on the construction of knowledge flow network. Among the patent-related data, citation information is known to be useful in depicting how knowledge flows across diverse technology classes [20]–[23] because it has the ability to capture the complex relationships between technology classes by clarifying technological antecedents and descendants [24]. In this regard, the patent citation information has been used in a variety of studies for the analysis of technological spillover effects including measuring the extent of the spillover effects according to four knowledge flow patterns [5], identifying opportunities for new technology development [25], investigating technological innovation capabilities in Africa by manifesting various types of knowledge spillovers [26], and examining the influence of patent citations on firms growth forecasts [27]. However, the latest patents do not have enough time to be cited, so the citation-based analysis has a limitation in that it does not reflect recent technological trends properly [6]. To address this limitation, patent co-classification analysis has been widely adopted for the identification of knowledge flows among technology areas since it can display direct knowledge flows by extracting the IPCs assigned to each patent [28], [29]. Several studies have combined network analysis methods with the co-classification analysis to explore the technological spillover effects from these direct influence relationships [12], [29], [30]. Park and Yoon [29] apply the Social Network Analysis (SNA) to the directed knowledge flow network constructed through the co-classification analysis. They evaluate the intermediarity of technology classes by calculating betweenness centrality and try to explain the technological spillover effects based on the measured intermediarity. However, the intermediarity does not adequately explain the spillover effects since the centrality shows only the degree to which a class is directly or indirectly close to other classes. Similarly, Lim and Park [30] investigate which roles of intermediaries the technology classes perform in the relative industry in the knowledge flow network generated by the co-classification analysis

TABLE 1. Three measures to determine the interestingness of association rule $A \rightarrow B$.

measure	Formula	Description
Support	$P(A \cap B)$	It examines the rule's usefulness. A rule with a high support value implies that the co-occurrence of the antecedent and consequent items in the rule is relatively frequent.
Confidence	$P(B\|A)$	It examines the rule's certainty. A rule with a high confidence value implies that the antecedent item is highly associated with the consequent item.
Lift	$\frac{P(B\|A)}{P(B)}$	It examines the correlation between the antecedent and consequent items. A lift greater than 1 means a positive correlation and a rule with a high lift value implies that the co-existence of the items is more likely not just a random occurrence, but rather due to the implication relationships.

including intra-industry mediator, inter-industry mediator, outward diffuser, and inward absorber. However, they only identify the roles of technology classes and do not quantify the extent of the spillover effects that each class has on other classes. Kim *et al.* [12] evaluate technological cross-impacts and identify core technologies using Analytic Network Process (ANP). However, ANP only considers the extent to which a class affects others. To measure the comprehensive spillover effects, the extent to which a class is affected by others should also be considered.

There have been, of course, lots of studies related to the technological spillover effect analysis that are not based on the patent data such as analyzing innovation intermediaries within industrial clusters' knowledge systems [31], investigating factors that negatively affect the spillover effect in reaction to foreign direct investment [32], and analyzing the impact of international channels on spillover for technological advance [33]. These studies, however, have mainly focused on the identification of influential factors and the extent to which they affect the knowledge spillovers from a particular industry perspective. However, for the applicability in the KOTEC, the extent of the technological spillover effects on all the technology classes should be quantifiable and measurable. The patent analysis can be an excellent tool to make this possible, so this study is based on it. Most patent analysis-based studies generally utilize bibliometric data of patent documents including patent classification codes and citation information [24], [34]. Using only bibliometric data naturally tends to exclude the technological implications described in patent documents [5], [35]. To remedy this problem, lots of studies have tried to encompass the technological descriptions into the patent analysis processes by using text mining techniques [36]. However, this study aims to measure the technological spillover effects in the perspective of technology classes (not individual technology), so this study uses only the bibliometric data of patents.

B. ASSOCIATION RULE MINING

Association rule mining is an unsupervised learning technique for discovering significant relationships between items in a given database [37]–[40] under an assumption that if people frequently purchase two items together in a single transaction, there must be a hidden relation between the co-purchased items [41]–[43]. The relationship is formulated as a form of a rule like $A \rightarrow B$ where A is the antecedent and B is the consequent which means that who purchases the item A will also generally tend to buy the item B [44], [45]. A number of studies have applied the association rule mining technique to identify core technology classes analyzing technological cross-impact [12], present an approach to design convergent product concepts [37], handle textual data for industrial knowledge management [46], propose a probabilistic model that can improve the efficiency of detecting medication errors [47], and present a stock market portfolio recommender system [48].

Three measures, support, confidence, and lift, are examined to determine the interestingness of generated association rules as shown in Table 1 [17], [38], [49], [50]. The support measure formulated as $P(A \cap B)$ indicates the probability that items A and B occur simultaneously in all transactions. The confidence measure formulated as $P(B|A)$ denotes the conditional probability that the consequent item B of the association rule occur in transactions given that the antecedent item A has already occurred in the same transaction. The lift measure formulated by dividing the confidence value by the probability of the consequent item B shows the statistical correlation between items A and B. The Apriori algorithm [51] is the most representative technique to generate association rules that make uses of pre-defined minimum threshold values of support, confidence, and lift [52]. It first collects frequent itemsets that have higher support values than the minimum threshold value, and then generate rules using the collected itemsets that have confidence and lift values exceeding the corresponding threshold values.

C. DECISION MAKING TRIAL AND EVALUATION LABORATORY

DEMATEL investigates the interrelationship structure among the relative factors to make decisions in complex problems [53], [54]. It measures the extent of direct and indirect effects of the factors by capturing their directed and weighted relationships and then makes the priorities by quantifying the impact and causality [55]. The procedure for applying DEMATEL is as follows [56] : 1) constructing a direct relation matrix which shows the direct effects between factors, 2) normalizing the direct relation matrix by dividing all elements by the maximum value of the sum of rows and columns so that the value of each element falls into between 0 and 1, and 3) computing a total relation matrix using convergent solutions to present comprehensive causal relationships by integrating the direct and indirect effects. In this study, DEMATEL is used to assess the extent of technological

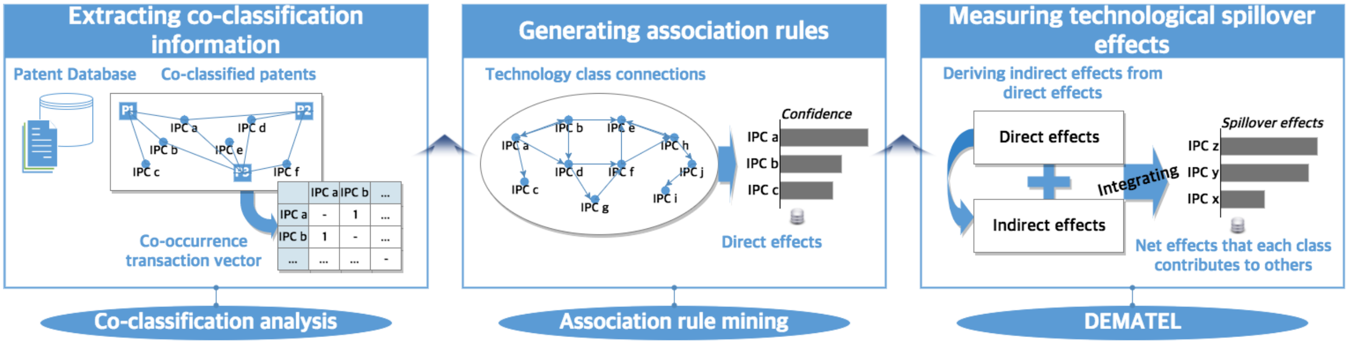


FIGURE 1. **Procedural framework for technological spillover effect analysis approach.**

spillover effects that represents the net effects that each technology class contributes to others. Various studies have used DEMATEL to examine the influential effects of factors such as generating technology impact networks [57], exploring the effects of technological knowledge spillovers [14], identifying a set of criteria for evaluating a green project management combining DEMATEL and analytical network process [58], and investigating firms innovation capability evaluation factors [59].

III. TECHNOLOGICAL SPILLOVER EFFECT ANALYSIS APPROACH

To explore how to analyze technological spillover effects, we present a procedural framework which consists of 3 steps as shown in Fig. 1: 1) extracting co-classification information from patent data, 2) generating association rules that represent the direct effects between technology classes, and 3) measuring technological spillover effects by deriving the indirect effects from the direct ones and then integrating the both effects.

A. EXTRACTING CO-CLASSIFICATION INFORMATION

A patent may generally be classified into multiple IPCs to indicate that the inventive solution implied in the patent has applicability in diverse technology domains. The fact that technological knowledge has been claimed in various technology fields means that this knowledge can be utilize in these fields, so capturing this point enables us to derive the aspect of knowledge sharing and transfer between technology areas [29]. In this regard, the patent co-classification analysis depicts the extent of technological knowledge exchange among technology classes by extracting co-classified information from patents. The IPC is represented as a set of alphanumeric codes and consists of hierarchical sets of sections, classes, subclasses and groups. In general, most studies related to the technology trend analysis based on the patent data have used the IPC subclasses that allows for the creation of an appropriate number of technology classes with clear technological boundaries [13], [15], [29], [60], so this study also utilizes the IPC subclasses.

B. GENERATING ASSOCIATION RULES

The association rule mining generates meaningful relationships between items or itemsets using the information about the co-purchased items in the same transaction. This study attempts to identify connection rules using the information about the co-occurred IPCs in the same patent. Therefore, to apply the association rule mining approach to our study, we consider a patent document and co-occurred technology classes in the patent as a transaction and co-purchased items in the transaction, respectively. Based on this assumption, we can generate association rules by investigating how many times the technology classes appear simultaneously in the same patents. We examine three measures, support, confidence, and lift, to determine the interestingness of the generated association rules. The support measure means how often the technology classes in the rules occur in all the patents. If the support value is low, the usefulness of the rules is lowered. Therefore, we pre-define the minimum threshold value and select only rules that have a support value above this one to ensure that only useful rules are applied in subsequent analysis. The lift measure examines the correlation between the antecedent and the consequent technology classes. If this value is less than 1, the negative correlation exists. Therefore, only rules with a lift value of 1 or more are selected to ensure that only rules with positive correlation can be used in subsequent analysis. The confidence measure means how closely the technology classes in the rules are associated. In this study, we use this confidence value as the extent of direct influence of the antecedent class on the consequent class because it shows how strong the association between the two classes is.

C. MEASURING TECHNOLOGICAL SPILLOVER EFFECTS

The DEMATEL comprehensively assesses the degree to which each relative factor affects and is affected in a complex network. To apply the DEMATEL, it is necessary to measure the direct influence relationships between relative factors and construct a direct relation matrix based on the measurement result. In this study, we establish a direct relation matrix using the association rules generated in the previous step since

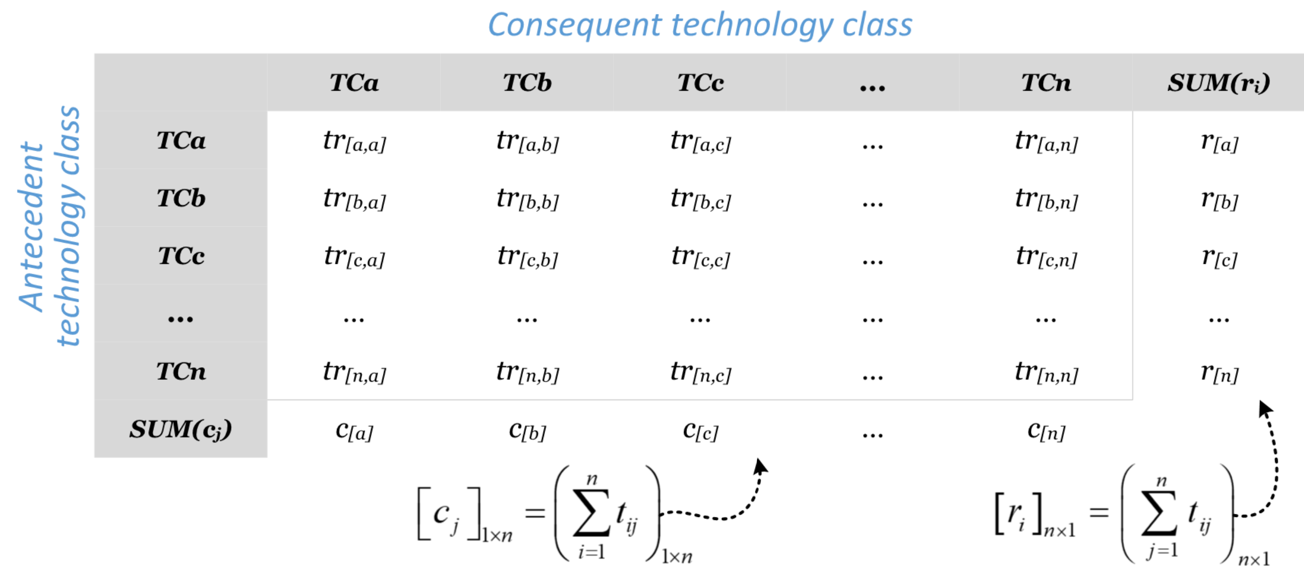


FIGURE 2. Total relation matrix.

TABLE 2. Summary of descriptive statistics and correlation coefficients.

	By IPC subclass (n = 615)				By IPC subclass (n = 615)			
	Mean	Stdev.	1	2	Mean	Stdev.	1	2
1. All available patents	53,406	32,744	1	-	695	1,763	1	-
2. Refined patents	20,664	10,338	0.889[a]	1	269	561	0.832[a]	1

these rules show direct influential relationships between technology classes. After normalizing the direct relation matrix by dividing all elements by the maximum value of the sum of rows and columns, we create a total relation matrix that represents the overall comprehensive spillover effects including direct and indirect influence relationships. In the total relation matrix, each cell value indicates the degree to which antecedent class directly or indirectly affects the consequent class, so the sum of the rows represents the degree to which each class affects all other classes and the sum of the columns means the degree of influence that each class receives from all others (Fig. 2). We name the sum of the rows as cause and the sum of the columns as effect. The sum of the cause and the effect indicates the impact of how much influence the class has on the technological spillover network and the difference of the two implies the causality that shows whether the class mainly influences or is affected.

IV. ILLUSTRATION

A. CO-CLASSIFICATION INFORMATION EXTRACTION

To extract the patent co-classification information, we collect patents granted in the KIPO. The total number of the collected patents is 427,250. Among them, we only use the patents that classified into two or more IPC subclasses since it is unnecessary to have patents classified into a single IPC in the analysis of technological spillover effects between technology classes. Form this constraint, 165,311 patents remain and these refined patents will be only considered for the technological spillover effect analysis. In this sense, we use only about 39% of all available patent data. It indicates that we should check whether this study lose the generality or not due to the use of only a few data sets. To do this, we compare the distribution of all available patents and that of the refined patents by IPC by measuring the Pearsons correlation coefficients. As shown in Table 2, the correlation coefficients are statistically significant at the 0.01 level. Note that, the number of the refined patents is always smaller than the number of all available patents in all the IPC sections and subclasses. Thus, the correlation coefficients naturally show high positive values. Nevertheless, the fact that the coefficients are greater than 0.8 indicates that there are very strong positive linear relationships between them. Therefore, although a few patent data sets are used, this study can be thought of as not losing the generality because the refined patents can adequately represent all available patents. Using the refined patent data, we calculate the co-occurrence frequency of pairs of technology classes as shown in Table 3. This co-occurrence information will be used in the next step as input data to generate association rules between technology classes.

B. ASSOCIATION RULE GENERATION

To generate association rules using the Apriori algorithm, it is required to define threshold values for the three measures, support, confidence, and lift. First, we specify a minimum

TABLE 3. Co-occurrence frequency of pairs of technology classes (top 20).

Pair of technology classes		Co-occurrence frequency	Pair of technology classes		Co-occurrence frequency
A61K	A61P	5,222	C08J	C08L	1,373
H04B	H04W	4,231	G06F	H04L	1,255
H04B	H04L	3,183	G01R	H01L	1,185
H04L	H04W	2,713	C12N	C12Q	1,165
C08K	C08L	2,251	G06F	G06Q	1,116
A61K	A61Q	2,228	G02F	G09G	1,015
B01D	C02F	1,668	C21D	C22C	1,002
F21S	F21V	1,511	G02F	H01L	1,002
G02B	G02F	1,437	A61P	C07D	981
A61K	C07D	1,430	H01L	H05B	961

TABLE 4. Generated association rules (top 20).

Antecedent class	Frequency (support)	Consequent class	Frequency (support)	Confidence
A61Q	2,589 (1.57%)	A61K	12,534 (7.58%)	86.06%
A01P	520 (0.31%)	A01N	1,135 (0.69%)	80.00%
G06N	111 (0.07%)	G06F	9,156 (5.54%)	69.37%
A01H	457 (0.28%)	C12N	5,825 (3.52%)	67.83%
C10N	96 (0.06%)	C10M	249 (0.15%)	67.71%
A23P	433 (0.26%)	A23L	3,379 (2.04%)	67.44%
A61P	7,964 (4.82%)	A61K	12,534 (7.58%)	65.57%
F25C	193 (0.12%)	F25D	843 (0.51%)	63.21%
F21Y	1,548 (0.94%)	F21V	4,380 (2.65%)	60.72%
A47H	83 (0.05%)	E06B	1,367 (0.83%)	60.24%
A43C	83 (0.05%)	A43B	316 (0.19%)	60.24%
A23F	214 (0.13%)	A23L	3,379 (2.04%)	58.41%
B22C	156 (0.09%)	B22D	1,051 (0.64%)	57.69%
F21W	671 (0.41%)	F21V	4,380 (2.65%)	56.33%
A45F	175 (0.11%)	A45C	532 (0.32%)	56.00%
C12H	193 (0.12%)	C12G	514 (0.31%)	54.92%
H04H	401 (0.24%)	H04N	5961 (3.60%)	54.11%
B29L	93 (0.06%)	B29C	3,415 (2.07%)	53.76%
A23G	302 (0.18%)	A23L	3,379 (2.04%)	53.64%
C12G	514 (0.31%)	C12R	1,597 (0.97%)	53.31%

threshold value for the support measure that examines the usefulness of the generated rules. If the threshold value is low, a large number of rules will be generated including technology classes having very low appearance frequencies. To set the minimum threshold value for the support measure, we extract the co-occurrence frequencies of the technology classes from the collected patent data and derive the association rules directly from them. And then, we calculate the average support value of all derived rules to determine the minimum threshold value based on the average value. The average support value is 0.36% so we set the minimum threshold value to 0.05% which is lower than the average value to ensure that the technology classes with extremely low frequencies of occurrence are not included in the final association rule set, but a sufficient number of rules can be obtained for subsequent analysis. Second, we determine a minimum threshold value for the confidence measure that examines the certainty of the rules. If the threshold value is low, rules with very low certainties will also be generated. The minimum threshold value determines whether a slight influence relationship between technology classes is taken into account or not. The confidence values of the association rules generated in this step are judged as the quantitative level of the direct influence relationships and the comprehensive spillover effects will be calculated based on these direct relationships at the subsequent step. It implies that the rules having slight direct relationships will not be reflected much in the spillover effects. Therefore, we set the minimum threshold value for the confidence measure to 0.05% just as the minimum support threshold value. Finally, we set a threshold value for the lift measure that examines the correlation between the antecedent and consequent technology classes. If this value is less than 1, it means that there is a negative correlation between them. Therefore, we set a constraint that the lift value must be greater than 1 to ensure that only rules with positive correlation can be used in subsequent analysis. Applying the Apriori algorithm based on these constraints leads to generate 33,638 association rules (Table 4).

TABLE 5. Total relation matrix of technology classes.

	A01B	A01C	A01D	A01F	A01G	A01H	A01K	A01M	A01N	A01P
A01B	0.1455	0.2578	0.2818	0.0540	0.4537	0.0457	0.1193	0.0739	0.1277	0.0649
A01C	0.2278	0.1059	0.0776	0.0204	0.5316	0.0561	0.1332	0.0985	0.1801	0.0923
A01D	0.3405	0.1062	0.1139	0.1102	0.4055	0.0445	0.1338	0.0540	0.1190	0.0602
A01F	0.1279	0.0544	0.2163	0.0431	0.2866	0.0512	0.1199	0.0442	0.1248	0.0578
A01G	0.0524	0.0698	0.0385	0.0137	0.3084	0.0538	0.1321	0.0807	0.1338	0.0614
A01H	0.0251	0.0350	0.0201	0.0120	0.2653	0.1442	0.1201	0.0406	0.2379	0.1061
A01K	0.0305	0.0385	0.0284	0.0130	0.2992	0.0551	0.1086	0.0606	0.1310	0.0614
A01M	0.0466	0.0704	0.0280	0.0118	0.4415	0.0466	0.1482	0.0658	0.1698	0.0844
A01N	0.0288	0.0466	0.0220	0.0119	0.2683	0.0959	0.1159	0.0617	0.5946	0.5994
A01P	0.0325	0.0528	0.0247	0.0122	0.2715	0.0940	0.1199	0.0676	1.3097	0.4951

C. TECHNOLOGICAL SPILLOVER EFFECT MEASUREMENT

We first construct an initial direct relation matrix using the relationships and the confidence values represented by the association rules generated in the previous step. With this as a starting point, this step applies the DEMATEL method to finally produce technological spillover effects. The normalized direct relation matrix can be obtained by dividing all the cell values of the direct relation matrix by the maximum value between the sum of the rows and the sum of the columns. In this case study, the maximum value is 9.4206. And then, the total relation matrix is computed by using convergent solutions to present comprehensive causal relationships by integrating the direct and indirect effects (Table 5). In the total relation matrix, the sum of the rows expresses the cause, the sum of the columns represents the effect, and the sum of the cause and the effect indicates the impact. We regard this impact value as a technological spillover effect of each technology class. However, as can be seen in Table 6 the effect value has a relatively larger variance than the cause value so summing them up simply makes the impact value largely biased towards the effect value. To remedy this problem, we normalize each of these and add them together. Table 7 shows the technological spillover effects for each technology class calculated using this approach.

TABLE 6. Summary of descriptive statistics of cause and effect.

Category	Min.	Max.	Average	Stdev.
Cause(r)	16.14	52.58	48.45	3.46
Effect(c)	1.36	656.35	48.45	77.30

H01L (semiconductor devices) appears to have the largest spillover effect value. It is believed that this is due to the special situation of Korea which has achieved rapid economic growth based on the electronics industry including the semiconductor business. From a global viewpoint, semiconductor-related technologies have had a significant impact on the development of advanced capabilities in a variety of industry fields including telecommunications, automobiles, and consumer electronics [61]. These technological trends seem to be reflected in the results of the technological spillover effect analysis. A61K (preparations for medical, dental, or toilet purposes) is mainly about the prevention or alleviation of abnormal conditions of the living body so it can be a main technological application domain where technology development occurs by absorbing the technological knowledge from the external areas including organic chemistry. This class has innovative impact on various medical-related fields such as nanobiochips and nanomedicine through the convergence with the diagnostic and therapeutic classes [15]. H04B (transmission) is largely related to the transmission systems of measured values, control signals, and digital information. It has provided technological advances on the data-driven business analytics-related fields through the technological convergence with the information technology and data science. It has played a pivotal role in technological convergence with multiplex communication, electric digital data processing, wireless communication networks, and broadcast and pictorial communication in information and communication technology standards [62].

V. DISCUSSION

A. SPILLOVER EFFECTS BY FINANCIAL FACTORS

This study focuses on the spillover effects from a technical point of view so we use only the patent data to measure the effects. However, in the spillover effects from an industrial point of view, besides technical aspects, financial factors can also have a significant impact. Therefore, it is also meaningful to understand whether the results of this study reflect the effects of financial factors. This study used patent data granted in the KIPO during 2010 and 2013. Just before this time, there was a financial crisis triggered by the Lehman Brothers collapse and Korea was also directly affected by the crisis. The analysis period of this study corresponds to the period of recovery from this crisis. To determine whether the spillover effects obtained from this study are consistent with this economic recovery trend, we calculate annual spillover effects (Table 8). Both the total amount and the average value of the spillover effects appear to largely reflect the recovery trend. It is, therefore, reasonable to assume that external factors such as financial crises have some influence on the spillover effects. It is obvious that quantitative assessment of

TABLE 7. Technological spillover effects between technology classes (top 20).

Technology class	Cause(r)	Effect(c)	Impact (r+c)	Normalized cause(n_r)	Normalized effect(n_c)	Spillover effect (n_r+n_c)
H01L	48.25	656.35	704.60	0.8811	1.0000	1.8811
A61K	51.94	556.86	608.80	0.9826	0.8481	1.8307
H04B	49.53	586.24	635.77	0.9164	0.8930	1.8094
H04W	49.94	438.13	488.07	0.9276	0.6668	1.5944
H04L	49.68	422.69	472.37	0.9203	0.6433	1.5636
A61P	52.24	354.33	406.57	0.9906	0.5389	1.5295
G06F	48.39	401.66	450.05	0.8850	0.6112	1.4962
C08L	47.41	357.49	404.90	0.8582	0.5437	1.4019
G02F	48.81	306.18	354.99	0.8967	0.4654	1.3621
C12N	51.15	258.97	310.12	0.9608	0.3933	1.3541
G01N	48.33	302.03	350.36	0.8835	0.4590	1.3425
B01D	48.01	296.47	344.48	0.8747	0.4506	1.3253
H04N	49.18	261.93	311.11	0.9067	0.3978	1.3045
G02B	48.68	258.20	306.88	0.8931	0.3921	1.2852
C08J	47.87	242.59	290.46	0.8708	0.3683	1.2391
C02F	48.35	227.44	275.79	0.8839	0.3452	1.2291
C07D	51.72	163.51	215.23	0.9766	0.2476	1.2242
C08K	47.68	218.31	265.99	0.8656	0.3312	1.1968
F21V	49.05	192.41	241.46	0.9032	0.2917	1.1949
A23L	50.69	151.64	202.33	0.9481	0.2294	1.1775

TABLE 8. Technological spillover effects by year.

Year	2010	2011	2012	2013
Sum of spillover effects	394.25	431.82	450.78	442.90
Average of spillover effects	0.9042	0.9596	0.9907	0.9670

the degree of influence by financial factors is very important to precisely measure the comprehensive spillover effects of a combination of technical and financial aspects. To do this, it is necessary to utilize a wide range of financial data including R&D investment, sales, operating profits, and the number of employees in each industry. Although investigating financial spillover effects caused by some external factors is crucial, it is beyond the scope of this study focusing on the technological spillover effects, so we will leave it as a future study.

B. COMPARISON BETWEEN PROPOSED APPROACH AND CURRENT APPROACH

This study proposed a systematic approach to measure the technological spillover effects between technology classes. To examine the feasibility of the proposed approach, we investigate how the measurement results of the approach explain the default probability of the obligors who have been guaranteed by the KOTEC. It is important to show that the beneficiaries of the technology finance system achieve great economic growth for the system utilizing the technology-based credit guarantees to be successful. The KOTECs greatest risk in continuing with the technology-based credit guarantee system is that the beneficiaries are in default. Therefore, this study attempts to explore the feasibility of the proposed approach by examining how the spillover effects explain the probability of default. To do that, we assign a rating for each technology class according to the corresponding spillover effects and collect the results of technology evaluation of several firms performed by the KOTEC. Some of those firms are now in default. And then, we construct a logit model which is formulated as:

$$\operatorname{logit}(p) = \log \frac{p}{1-p} = \alpha - \beta x \tag{1}$$

where p denotes the probability of default and x means the score in accordance with the spillover effect rating. We conduct a ROC curve analysis to determine the extent to which the spillover effects explain the probability of default. The total number of technology evaluation cases used in the feasibility verification process is 7,862. Table 9 summarizes the constructed logit models based on the proposed approach and the KTRSs current evaluation approach for the spillover effects. First, the proposed approach shows an increase in default rates as spillover effect ratings are lowers whereas the current approach has little or no correlation between the default rates and the ratings. It implies that the proposed approach better explains the relationship between the spillover effects and the probability of default than the current approach. Second, the logit model generated through the proposed approach is statistically significant at the 0.01 level, but no significant model is created through the current approach. A positive β means that the higher the spillover effect rating, the lower the logit for the default probability. Finally, the proposed approach offers an improvement over

TABLE 9. Comparison of logit models of proposed approach and current approach.

Approach	Default rates per spillover effect rating				β	p-value	AUC
	A	B	C	D			
Current approach	6.74%	6.34%	7.32%	4.02%	-0.042	0.614	0.500 (default)
Proposed approach	5.72%	7.12%	8.75%	9.68%	0.207	0.010	0.551

TABLE 10. Comparison of logit models of proposed approach and current approach.

Approach	Recall				Precsion			
	DT	RF	MLP	Average	DT	RF	MLP	Average
Proposed approach	0.8336	0.8277	0.6709	0.7774	0.9129	0.9097	0.5283	0.7836
SNA-based approach [29]	0.8365	0.8412	0.4520	0.7099	0.8742	0.8716	0.5264	0.7574
ANP-based approach [12]	0.7611	0.8180	0.6491	0.7427	0.9109	0.8843	0.5471	0.7808

the current approach with an AUC value of 0.551. Of course, the AUC value is not absolutely high. Our approach is based solely on the patent data, but the default probability is basically related to financial performance. It indicates that it is impossible for our approach to accurately predict whether a firm will be in default or not. Moreover, the purpose of our study is not to present a classification model, but to suggest an approach to measure the technological spillover effects in the perspective of technology classes. We only want to compare our approach with the current approach using the AUC values. In this sense, we can conclude that the approach proposed in this study has better performance in several aspects than the spillover effect evaluation model currently used in the KOTEC.

C. COMPARISON BETWEEN PROPOSED APPROACH AND OTHER RELEVANT APPROACHES

To determine the performance of our approach, we have created additional models for predicting the probability of default using other relevant approaches, the SNA-based approach [29] and the ANP-based approach [12]. Recall and precision metrics are used to measure the predictive performance of these models. If we set true when default occurs and false when default does not occur, then the recall and precision metrics can be formulated as:

$$\mathrm{RECALL} = \frac{TP}{TP + FN}, \quad \mathrm{PRECISION} = \frac{TP}{TP + FP} \tag{2}$$

where TP (True Positive) denotes the number of instances which are labeled as positive and are also detected as positive, FN (False Negative) denotes the number of instances which are labeled as positive but are wrongly detected as negative, and FP (False Positive) denotes the number of instances which are labeled ad negative but are wrongly detected as positive [63]. Recall and precision metrics have been widely used to measure the performance of classification models when sample data is imbalanced [64]. The dataset used in this comparison is so imbalanced that the ratio of the default cases is only 2.8%. Therefore, we use the Synthetic Minority Over-sampling Technique (SMOTE) to improve the classification performance for the imbalanced dataset. The SMOTE is a technique for generating a balanced dataset by oversampling the data samples using statistical distributions [65].

This analysis used Python and the machine learning library scikit-learn to create the classification models, and imbalanced-learn for applying the SMOTE. The total number of dataset has been increased to 13,162 by oversampling so that the number of default cases was equal to the number of normal cases. To generate classification models, we used Decision Tree (DT), Random Forest (RF), and Multi-Layer Perceptron (MLP), which are widely used in the relevant problems. Hyperparameter tuning was not employed in this analysis for general performance comparison. In the case of MLP, we create 3 hidden layers and 1,024, 512, and 8 hidden nodes in each layer. Table 10 shows the performance comparison results of the classification models generated by the approach proposed in this study and the relevant previous approaches. On the whole, the proposed approach appears to have a slightly better recall and precision performance than previous approaches. In this sense, we can also conclude that the proposed approach is relatively meaningful in that it can reasonably predict the probability of default.

VI. CONCLUSION

For a successful implementation of the technology financing system, it is required to establish a clear technology evaluation system. Technology evaluation institutes including KOTEC have generally relied on the manual work by relevant experts for the technology evaluation. In order to improve the efficiency of this manual work, various quantitative indicators for the assessment of the potential future value of technology have been used together. These indicators are primarily aimed at evaluating technology classes, not individual technology. Among these indicators, the spillover effect is generally perceived as useful for the disposal of patents of a firm which received credit guarantee but is in default. A model for measuring the spillover effects of technology classes is currently presented in the KTRS, but it has low reliability. Therefore, this study proposed a systematic approach to measuring technological spillover effects between technology classes. To measure the extent to which technology classes affect and are affected by other classes on the technological knowledge flow network, we first extracted co-classification information from patent data, and then generated association rules between technology classes employing the association rule mining. The relationships represented by the rules, however, can only depict the direct connections between technology classes, so we finally utilized the DEMATEL method to obtain comprehensive spillover effects including direct and indirect influential relationships. We expect that this study can contribute to establish a quantitative evaluation model to help assess technologies for successful technology-based credit guarantee system. It will improve the reliability of the technology assessment by reducing the variance of the qualitative evaluation results due to the individual differences of the evaluator. Furthermore, it will also contribute to enhance the efficiency of evaluation work. Despite the contribution, further research problems still remain to be investigated. First, the AUC value of the proposed approach is superior to the current approach, but not absolutely high. The approach needs to be improved in a way that utilizes the patent data to reflect technological viewpoints as well as the financial data to reflect economic viewpoints. Second, the motivation of this study is to measure the spillover effects on the technology classes to help the relevant experts in evaluating technologies manually. It leads to the identical spillover effects of all individual technologies belonging to the same technology class. Therefore, to advance to a sophisticated model, how to measure the spillover effects on the individual technology should be discussed. Third, this study utilized only the bibliometric information of the patent data. We also need to utilize the invention descriptions in the patent documents to enhance the results of our analysis. Using text mining techniques can be a good way to do this. Fourth, the feasibility verification of the approach was made only to the probability of default. The verification needs to be conducted in terms of financial performance including sales and operating profits, which will make the proposed approach more robust. Finally, this study focused only on the technological spillover effects. However, to precisely measure the comprehensive spillover effects, it is necessary to investigate the influence by financial factors. We must exploit a wide variety of financial data to explore financial effects that reflect the financial factors.

ACKNOWLEDGMENT

W. Seo and M. Noh thank to the KOTEC for providing relevant data essential for this study.

REFERENCES

[1] L. I. Shangzhen, "A literature review of technology and finance," *Can. Social Sci.*, vol. 11, no. 11, pp. 89–93, 2015.

[2] D. T. V. Trang, "Experiences of green credit development—Lessons learned to Vietnam," *Rev. Bus. Econ. Stud.*, vol. 4, no. 1, pp. 85–91, 2016.

[3] J. Shin, "New business model creation through the triple helix of young entrepreneurs, SNSs, and smart devices," *Int. J. Technol. Manage.*, vol. 66, no. 4, pp. 302–318, 2014.

[4] Y.-J. Liu and S.-J. Liu, "The intellectual property policy of taiwan: A strategic viewpoint," in *Proc. IEEE Int. Eng. Manage. Conf.*, vol. 1. Oct. 2004, pp. 42–46.

[5] N. Ko, J. Yoon, and W. Seo, "Analyzing interdisciplinarity of technology fusion using knowledge flows of patents," *Expert Syst. Appl.*, vol. 41, no. 4, pp. 1955–1963, 2014.

[6] J. Yoon and K. Kim, "Identifying rapidly evolving technological trends for R&D planning using SAO-based semantic patent networks," *Scientometrics*, vol. 88, no. 1, pp. 213–228, 2011.

[7] W. Seo, J. Yoon, H. Park, B.-Y. Coh, J.-M. Lee, and O.-J. Kwon, "Product opportunity identification based on internal capabilities using text mining and association rule mining," *Technol. Forecast. Social Change*, vol. 105, pp. 94–104, Apr. 2016.

[8] H. Fujii, K. Yoshida, and K. Sugimura, "Research and development strategy in biological technologies: A patent data analysis of Japanese manufacturing firms," *Sustainability*, vol. 8, no. 4, pp. 351–365, 2016.

[9] J. Yoon, W. Seo, B.-Y. Coh, I. Song, and J.-M. Lee, "Identifying product opportunities using collaborative filtering-based patent analysis," *Comput. Ind. Eng.*, vol. 107, pp. 376–387, May 2017.

[10] A. Segev, S. Jung, and S. Choi, "Analysis of technology trends basedon diverse data sources," *IEEE Trans. Serv. Comput.*, vol. 8, no. 6, pp. 903–915, Nov. 2015.

[11] F. Yu and Y. Wu, "Patent citations and knowledge spillovers: An analysis of Chinese patents registered in the USA," *Asian J. Technol. Innov.*, vol. 22, no. 1, pp. 86–99, 2014.

[12] C. Kim, H. Lee, H. Seol, and C. Lee, "Identifying core technologies based on technological cross-impacts: An association rule mining (ARM) and analytic network process (ANP) approach," *Expert Syst. Appl.*, vol. 38, no. 10, pp. 12559–12564, 2011.

[13] S. Altuntas, T. Dereli, and A. Kusiak, "Analysis of patent documents with weighted association rules," *Technol. Forecast. Social Change*, vol. 92, pp. 249–262, Mar. 2015.

[14] J.-K. Chen and I.-S. Chen, "An inno-qual performance system for higher education," *Scientometrics*, vol. 93, no. 3, pp. 1119–1149, 2012.

[15] W. Seo, "Analyzing causality of technological knowledge spillovers: Patent analysis approach," *Procedia Manuf.*, vol. 2, pp. 485–489, Jan. 2015.

[16] Q. Deng, X. Liu, and H. Liao, "Identifying critical factors in the eco-efficiency of remanufacturing based on the fuzzy DEMATEL method," *Sustainability*, vol. 7, no. 11, pp. 15527–15547, 2015.

[17] S. Atapattu, C. Tellambura, and H. Jiang, "Analysis of area under the ROC curve of energy detection," *IEEE Trans. Wireless Commun.*, vol. 9, no. 3, pp. 1216–1225, Mar. 2010.

[18] T. Schoenecker and L. Swanson, "Indicators of firm technological capability: Validity and performance implications," *IEEE Trans. Eng. Manag.*, vol. 49, no. 1, pp. 36–44, Feb. 2002.

[19] S. Choi and H. Park, "Investigation of strategic changes using patent co-inventor network analysis: The case of samsung electronics," *Sustainability*, vol. 8, no. 12, pp. 1315–1327, 2016.

[20] K.-K. Lai and S.-J. Wu, "Using the patent co-citation approach to establish a new patent classification system," *Inf. Process. Manage.*, vol. 41, no. 2, pp. 313–330, 2005.

[21] H. Lee, C. Kim, H. Cho, and Y. Park, "An ANP-based technology network for identification of core technologies: A case of telecommunication technologies," *Expert Syst. Appl.*, vol. 36, no. 1, pp. 894–908, 2009.

[22] S. Lee, B. Yoon, and Y. Park, "An approach to discovering new technology opportunities: Keyword-based patent map approach," *Technovation*, vol. 29, no. 6, pp. 481–497, 2009.

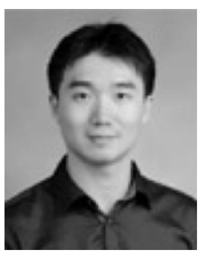
[23] T. E. Stuart and J. M. Podolny, "Local search and the evolution of technological capabilities," *Strategic Manage. J.*, vol. 17, no. 1, pp. 21–38, 1996.

[24] M. Trajtenberg, R. Henderson, and A. Jaffe, "University versus corporate patents: A window on the basicness of invention," *Econ. Innov. New Technol.*, vol. 5, no. 1, pp. 19–50, 1997.

[25] B. Yoon and Y. Park, "Development of new technology forecasting algorithm: Hybrid approach for morphology analysis and conjoint analysis of patent information," *IEEE Trans. Eng. Manage.*, vol. 54, no. 3, pp. 588–599, Aug. 2007.

[26] C.-Y. Tseng, "Technology development and knowledge spillover in Africa: Evidence using patent and citation data," *Int. J. Technol. Manage.*, vol. 45, nos. 1–2, pp. 50–61, 2008.

[27] J. Bae, W. Hur, J. Lee, and J. Goh, "Patent citations and financial analysts' long-term growth forecasts," *Sustainability*, vol. 9, no. 5, pp. 846–860, 2017.
[28] B. Verspagen, "Measuring intersectoral technology spillovers: Estimates from the European and US patent office databases," *Econ. Syst. Res.*, vol. 9, no. 1, pp. 47–65, 1997.
[29] H. Park and J. Yoon, "Assessing coreness and intermediarity of technology sectors using patent co-classification analysis: The case of Korean national R&D," *Scientometrics*, vol. 98, no. 2, pp. 853–890, 2014.
[30] H. Lim and Y. Park, "Identification of technological knowledge intermediaries," *Scientometrics*, vol. 84, no. 3, pp. 543–561, 2010.
[31] J. Guo and B. Guo, "How do innovation intermediaries facilitate knowledge spillovers within industrial clusters? A knowledge-processing perspective," *Asian J. Technol. Innov.*, vol. 21, no. 2, pp. 31–49, 2013.
[32] J. Du, H. Li, and X. Wu, "Empirical analysis on the negative technology spillover effect of foreign direct investment in China," *Asian J. Technol. Innov.*, vol. 16, no. 2, pp. 133–151, 2008.
[33] W. Fu and J. R. Díez, "Knowledge spillover and technological upgrading: The case of Guangdong province, China," *Asian J. Technol. Innov.*, vol. 18, no. 2, pp. 187–217, 2010.
[34] F. Narin, "Patent bibliometrics," *Scientometrics*, vol. 30, no. 1, pp. 147–155, 1994.
[35] S. Choi, H. Kim, J. Yoon, K. Kim, and J. Y. Lee, "An SAO-based text-mining approach for technology roadmapping using patent information," *R&D Manage.*, vol. 43, no. 1, pp. 52–74, 2013.
[36] S. Choi, H. Park, D. Kang, J. Y. Lee, and K. Kim, "An SAO-based text mining approach to building a technology tree for technology planning," *Expert Syst. Appl.*, vol. 39, no. 13, pp. 11443–11455, 2012.
[37] C. Lee, B. Song, and Y. Park, "Design of convergent product concepts based on functionality: An association rule mining and decision tree approach," *Expert Syst. Appl.*, vol. 39, no. 10, pp. 9534–9542, 2012.
[38] M. Song, G. E. Heo, and Y. Ding, "SemPathFinder: Semantic path analysis for discovering publicly unknown knowledge," *J. Informetrics*, vol. 9, no. 4, pp. 686–703, 2015.
[39] H.-N. Su, "Global interdependence of collaborative R&D-typology and association of international co-patenting," *Sustainability*, vol. 9, no. 4, pp. 541–568, 2017.
[40] T. P. Exarchos, C. Papaloukas, D. I. Fotiadis, and L. K. Michalis, "An association rule mining-based methodology for automated detection of ischemic ECG beats," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 8, pp. 1531–1540, Aug. 2006. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/16916087
[41] M.-J. Shih, D.-R. Liu, and M.-L. Hsu, "Discovering competitive intelligence by mining changes in patent trends," *Expert Syst. Appl.*, vol. 37, no. 4, pp. 2882–2890, 2010.
[42] R. J. Kuo, C. M. Chao, and Y. Chiu, "Application of particle swarm optimization to association rule mining," *Appl. Soft Comput.*, vol. 11, no. 1, pp. 326–336, 2011.
[43] D. Zhang and L. Zhou, "Discovering golden nuggets: Data mining in financial application," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 34, no. 4, pp. 513–522, Nov. 2004.
[44] R. J. Bayardo, Jr., R. Agrawal, and D. Gunopulos, "Constraint-based rule mining in large, dense databases," *Data Mining Knowl. Discovery*, vol. 4, nos. 2–3, pp. 217–240, 2000.
[45] R. Wolff and A. Schuster, "Association rule mining in peer-to-peer systems," *IEEE Trans. Syst. Man, Cybern. B, Cybern.*, vol. 34, no. 6, pp. 2426–2438, Dec. 2004. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/15619943
[46] N. Ur-Rahman and J. A. Harding, "Textual data mining for industrial knowledge management and text classification: A business oriented approach," *Expert Syst. Appl.*, vol. 39, no. 5, pp. 4729–4739, 2012.
[47] P. A. Nguyen *et al.*, "A probabilistic model for reducing medication errors," *PLoS ONE*, vol. 8, no. 12, p. e82401, 2013. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/24312659
[48] P. Paranjape-Voditel and U. Deshpande, "A stock market portfolio recommender system based on association rule mining," *Appl. Soft Comput.*, vol. 13, no. 2, pp. 1055–1063, 2013.
[49] A. L. Buczak and C. M. Gifford, "Fuzzy association rule mining for community crime pattern discovery," in *Proc. ACM SIGKDD Workshop Intell. Secur. Inform.*, 2010, pp. 2–11.
[50] V. S. Verykios, A. K. Elmagarmid, E. Bertino, Y. Saygin, and E. Dasseni, "Association rule hiding," *IEEE Trans. Knowl. Data Eng.*, vol. 16, no. 4, pp. 434–447, Apr. 2004.
[51] R. Agrawal, T. Imieliński, and A. Swami, "Mining association rules between sets of items in large databases," in *Proc. ACM SIGMOD Int. Conf. Manage. Data*, 1993, pp. 207–216.
[52] K. Alam, R. Ahmad, and K. Ko, "Enabling far-edge analytics: Performance profiling of frequent pattern mining algorithms," *IEEE Access*, vol. 5, pp. 8236–8249, 2017.
[53] A. Gabus and E. Fontela, "World problems, an invitation to further thought within the framework of DEMATEL," Battelle Geneva Res. Center, Geneva, Switzerland, Tech. Rep., 1972.
[54] C.-Y. Hung and W.-Y. Lee, "A proactive technology selection model for new technology: The case of 3D IC TSV," *Technol. Forecast. Social Change*, vol. 103, pp. 191–202, Feb. 2016.
[55] J.-I. Shieh, H.-H. Wu, and K.-K. Huang, "A DEMATEL method in identifying key success factors of hospital service quality," *Knowl.-Based Syst.*, vol. 23, no. 3, pp. 277–282, 2010.
[56] C.-L. Lin and G.-H. Tzeng, "A value-created system of science (technology) park by using DEMATEL," *Expert Syst. Appl.*, vol. 36, no. 6, pp. 9683–9697, 2009.
[57] S.-S. Ko, N. Ko, D. Kim, H. Park, and J. Yoon, "Analyzing technology impact networks for R&D planning using patents: Combined application of network approaches," *Scientometrics*, vol. 101, no. 1, pp. 917–936, 2014.
[58] Y.-C. Chou, C.-H. Yang, C.-H. Lu, V. T. Dang, and P.-A. Yang, "Building criteria for evaluating green project management: An integrated approach of DEMATEL and ANP," *Sustainability*, vol. 9, no. 5, pp. 740–756, 2017.
[59] D. Sumrit and P. Anuntavoranich, "Using DEMATEL method to analyze the causal relations on technological innovation capability evaluation factors in Thai technology-based firms," *Int. Trans. J. Eng., Manage., Appl. Sci. Technol.*, vol. 4, no. 2, pp. 81–103, 2013.
[60] S.-J. Shin and W. Seo, "Identifying new technology areas based on firm's internal capabilities," *J. Administ. Bus. Stud.*, vol. 3, no. 3, pp. 113–120, 2017, doi: 10.20474/jabs-3.3.1.
[61] P. Adams, R. Fontana, and F. Malerba, "The magnitude of innovation by demand in a sectoral system: The role of industrial users in semiconductors," *Res. Policy*, vol. 42, no. 1, pp. 1–14, 2013.
[62] E. J. Han and S. Y. Sohn, "Technological convergence in standards for information and communication technologies," *Technol. Forecast. Social Change*, vol. 106, pp. 1–10, May 2016.
[63] X. Zhu, H. Tao, Z. Wu, J. Cao, K. Kalish, and J. Kayne, *Fraud Prevention in Online Digital Advertising*. Cham, Switzerland: Springer, 2017.
[64] N. V. Chawla, "Data mining for imbalanced datasets: An overview," in *Data Mining and Knowledge Discovery Handbook*. Boston, MA, USA: Springer, 2009, pp. 875–886.
[65] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, no. 1, pp. 321–357, 2002.



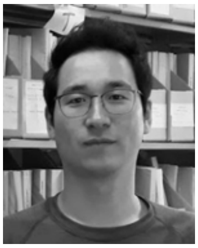
SUNGCHUL CHOI was born in Busan, South Korea, in 1982. He received the B.S. degree in management from Handong Global University in 2006 and the Ph.D. degree in industrial management engineering from POSTECH, South Korea, in 2012.

He was the Technology Strategy Manager with the CTO Office, Samsung Advanced Institute of Technology. He has conducted research on patent analysis, technology roadmapping, and strategic planning, using a text mining approach. He is currently an Assistant Professor with the Department of Industrial and Management Engineering, Gachon University. He is currently focusing on applying machine learning and deep learning approaches to patent big data and has been conducting various studies related to NLP and artificial intelligence.

Prof. Choi is a member of the Korean Institute of Industrial Engineers.

MAENG SEOK NOH was born in Busan, South Korea, in 1973. He received the B.S., M.S., and Ph.D. degrees from the Department of Statistics, Seoul National University, in 1996, 1998, and 2005, respectively. His thesis was on the analysis of binary data and robust modelling via hierarchical likelihood.

Since 2006, he has been a Professor with the Department of Statistics, Pukyong National University, Busan. His current research interests are application and software developments for hierarchical generalized linear models.



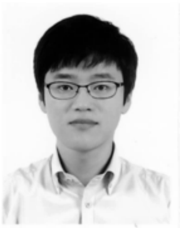
JANGHYEOK YOON was born in Daegu, South Korea, in 1979. He received the B.S., M.S., and Ph.D. degrees in industrial engineering from the Pohang University of Science Technology, Pohang, South Korea, in 2002, 2004, and 2011, respectively.

He was an IT Consultant and an Application Engineer for an affiliate of the LG Group from 2004 to 2007. He was an Associate Researcher with the Korea Institute of Intellectual Property, a Korean public research institute, from 2011 to 2012. Since 2012, he has been an Associate Professor with the Department of Industrial Engineering, Konkuk University, Seoul, South Korea. He has authored over 50 articles and holds ten patents related to patent mining-based technology intelligence and opportunity discovery systems. His research interests are in technology and service intelligence, which covers the analytical and multidisciplinary methodologies for technology, service (business model), and future planning, and thus they are based on the combined application of theories and analytics in various fields such as industrial engineering, business administration, and computer science.

Prof. Yoon is a member of the Korean Institute of Industrial Engineers and has been an Editor of the *Journal of Intellectual Property* since 2014.

HYUNSEOK PARK received the B.S. degree in media and communication engineering from Hanyang University, Seoul, in 2008, and the Ph.D. degree in technology and innovation management from the Pohang University of Science and Technology, Pohang, in 2014.

He joined the Fraunhofer Institute of Industrial Engineering as a Visiting Researcher in 2014. From 2014 to 2016, he was a Post-Doctoral Associate with the Institute for Data, Systems, and Society, Massachusetts Institute of Technology. Since 2016, he has been an Assistant Professor with the Department of Information Systems, Hanyang University. His research interests include technological change and developmental trajectories, quantification and identification of technological discontinuities, function-based networking of heterogeneous technologies, and patent-based analogical design.



WONCHUL SEO was born in Suncheon, South Korea, in 1980. He received the B.S. and Ph.D. degrees in industrial management engineering from the Pohang University of Science Technology, Pohang, South Korea, in 2003 and 2010, respectively.

He was a Senior Engineer with Samsung Electronics from 2010 to 2012 and an Associate Researcher with the Korea Institute of Intellectual Property, a Korean public research institute, from 2012 to 2013. Since 2013, he has been an Assistant Professor with the Division of Systems Management and Engineering, Graduate School of Management of Technology, Pukyong National University, Busan, South Korea. He has authored two books and over 40 articles. He holds one patent related to technological spillover effect analysis methodology. His research interests include patent analysis-based technology intelligence, technology and product opportunity discovery, technology trend identification by analyzing knowledge flow network, and quantification of technological spillover effects.